

УДК 303
ББК 60в6

ПРИМЕНЕНИЕ НЕЧЕТКОЙ ЛОГИКИ В ГУМАНИТАРНЫХ ИССЛЕДОВАНИЯХ

© М.И. Вершинин *, Л.П. Вершинина **, 2007

* Северо-Западный государственный заочный технический университет
191186, г. Санкт-Петербург, ул. Миллионная, д. 5** Санкт-Петербургский государственный университет культуры и искусств,
191186, г. Санкт-Петербург, Дворцовая набережная, д. 2

Многие человеческие знания и связи с внешним миром включают такие построения, которые нельзя назвать множествами в классическом смысле. Поэтому формальная логика не всегда приемлема для решения задач гуманитарных исследований, которые имеют дело не с количественными, а с качественными характеристиками. В статье показаны возможности и преимущества использования нечеткой логики для решения гуманитарных задач.

Ключевые слова: классическая математика, теория нечетких множеств, гуманитарные исследования.

В 1965 г. в журнале «Information and Control» появилась статья Л. Заде «Fuzzy sets» [1]. При переводе этого названия на другие языки возникло немало трудностей из-за семантической неоднозначности термина «fuzzy». На русский язык его переводили как «нечеткий», «размытый», «расплывчатый» и даже «неопределенный». Первый из вариантов со временем вытеснил все остальные и стал в отечественной литературе узаконенным.

Основная идея Л. Заде состоит в том, что человеческий способ рассуждения, опирающийся на естественный язык, не может быть описан в рамках традиционных математических формализмов. Человек мыслит нечеткими понятиями, делает заключения в неточных, неколичественных, нечетких терминах. Благодаря этой способности люди могут расшифровывать неразборчивый почерк, понимать искаженную речь, концентрировать внимание на той информации, которая приводит к решению.

Л. Заде явился основателем новой математической дисциплины, в основе которой лежит не классическая теория множеств, а теория нечетких множеств. Теория нечетких множеств – это, по сути дела, шаг на пути к сближению точности классической математики и всепроникающей нечеткости реального мира.

Множество или совокупность объектов – основное понятие в математике. Многие, возможно большинство, человеческих знаний и связей с внешним миром включают такие построения, которые нельзя назвать множествами в классическом смысле. Их скорее следует считать «нечеткими множествами» (или подмножествами), т.е. классами

с нечеткими границами, когда переход от принадлежности к классу к непринадлежности происходит постепенно, не резко. Логика человеческого рассуждения основывается не на классической или даже многозначной логике, а на логике с нечеткими значениями истинности, с нечеткими связками и нечеткими правилами вывода.

В классической математике есть только две приемлемые ситуации для элемента: он может либо быть, либо не быть элементом подмножества. Любая формальная логика, в том числе булева, основана на этом: элемент принадлежит или не принадлежит подмножеству данного множества. Понятие нечеткого подмножества основано на принадлежности элемента подмножеству в большей или меньшей степени.

С самого начала основная прагматическая цель Л. Заде – создание аппарата, способного моделировать человеческие рассуждения и объяснять человеческие приемы принятия решений в ходе решения различных задач, привлекла в эту область многочисленную армию прикладников. К потоку чисто математических работ добавился поток работ прикладных. Теория нечетких множеств нашла применение при создании систем, понимающих тексты на естественном языке, при создании планирующих систем, использующих неполную информацию, при управлении техническими, социальными и экономическими системами, в системах искусственного интеллекта и робототехнических системах. Теория нечетких множеств применяется при создании систем управления в технике, систем диагностирования в медицине, при разработке геологических и биологических моделей различных процессов и явлений.

Теория нечетких множеств позволяет наилучшим образом структурировать все то, что разделено не очень точными границами, например мысль, язык и восприятие у людей. Гуманитарные науки наполнены различными видами абстрактных форм: при исследовании гуманитарного объекта имеющуюся информацию очень часто нельзя формализовать на основе бинарной логики.

Источником расплывчатой неопределенности в гуманитарных исследованиях является неполное знание об объекте исследования, т.е. неполнота априорной информации, неполное обладание объекта некоторым свойством, неточность оценки объекта, субъективность критериев выбора, субъективность оценки и т.д.

Как правило, в гуманитарных исследованиях мы редко имеем дело только с количественными данными. Обычно задачи гуманитарных исследований – это задачи качественно-количественного анализа [2]. Что касается количественных свойств гуманитарных объектов, то с ними дело обстоит сравнительно просто. Для них имеются общепринятые эталоны измерения: один год для возраста или образовательной подготовки, один человек для исчисления размера семьи и т.д.

Качественные характеристики (социальная принадлежность, мнения людей и т.д.), как правило, не имеют установленных эталонов измерения [3]. Их приходится конструировать в соответствии с природой изучаемого объекта и согласно гипотезам исследования. Поэтому в целом измерительные процедуры качественных характеристик – своего рода изобретения, которым присуща субъективность.

Часто для измерения качественных свойств гуманитарных объектов не существует общепринятых шкал. Отсутствие объективной физической шкалы вносит неопределенность в процесс принятия решений. Рациональную шкалу для измерения таких свойств можно получить, применяя, например, основанную на нечеткой логике процедуру Саати [4].

В реальных ситуациях принятия решений в большей части субъективны и неточно определены также цели, ограничения, критерии выбора. Поэтому при построении моделей принятия решений в гуманитарных исследованиях возникает необходимость использования нечеткой логики. Задача достижения нечетко поставленной цели при нечетких ограничениях может быть решена на основе принципа слияния Беллмана–Заде [5]. Нечеткая цель и нечеткие ограничения описываются нечеткими подмножествами универсального множества альтернатив. При этом нечеткое решение определяется как нечеткое подмножество, получающееся в результате слияния нечетких целей и нечетких ограничений.

Особый интерес представляют попытки применения к задачам принятия решений теории возможности [6]. Нечеткая оценка возможности, понимаемая как субъективное отражение внутренних ограничений объекта, требует меньшего уровня априорной информированности, чем распределение вероятностей.

Использование стандартных количественных методов и статистических показателей взаимосвязи в гуманитарных исследованиях часто не вполне обоснованно создает впечатление математической точности выводов. Однако «жесткие» методы, как правило, не учитывают неопределенность и неоднозначность свойств изучаемых объектов, что ведет к упрощению, а то и к искажению реальности. Нередко на практике точные математические методы применяются к неточным моделям, где, например, нет никаких оснований задавать коэффициенты в виде точно определенных чисел. Такое искусственное сужение априорной информации может привести к искажению конечных результатов. Во избежание этого следует вводить поля допусков (интервалы толерантности) на коэффициенты модели, а затем использовать для решения задачи теорию нечетких множеств.

При решении ряда практических задач иногда нет необходимости в построении точных моделей, например в социологии и психологии при изучении тенденции, динамики социального процесса. Теория нечетких множеств дает возможность работать с качественно или неполно определенной информацией, рассматривать явления и процессы в многообразии их связей и в динамике, формируя при этом адекватные модели.

Формализация нечетких понятий и отношений обеспечивается введением лингвистической и нечеткой переменных, нечетких множества и отношения [7].

Лингвистическая переменная характеризуется нечисловыми значениями (термами). Например, в социологических исследованиях при делении обследуемого контингента на возрастные группы понятие «возраст» может оцениваться значениями «младший», «средний», «старший».

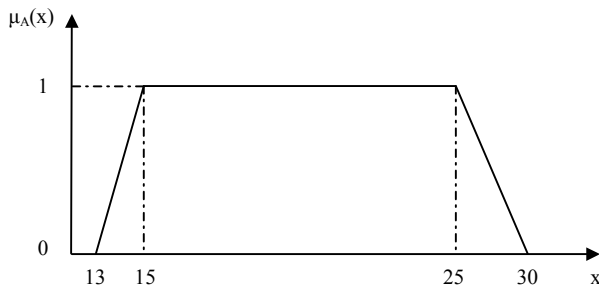
Нечеткая переменная характеризуется функцией принадлежности $\mu(x)$, принимающей значение из интервала $[0,1]$ в ее области определения X . Связь между значениями лингвистической и нечеткой переменной поясним следующим примером. С лингвистической переменной «возраст» V со значениями $V = \{\text{младший, средний, старший}\}$ сопоставляются соответственно три нечетких переменных со значениями (в годах) $V_{\text{мл}} \leq 30$; $25 \leq V_{\text{ср}} \leq 35$; $V_{\text{ст}} \geq 50$.

Функция принадлежности $\mu(x)$ принимает два крайних значения: 0 (отсутствие некоторого свойства) и 1 (полное обладание свойством) и проме-

жуточные: $0 < \mu(x) < 1$ (различная степень обладания свойством). Именно диапазон $0 < \mu(x) < 1$ характеризует нечеткость переменной. В примере нечеткость переменных $V_{мл}$, $V_{ср}$, $V_{ст}$ характеризуется диапазоном значений 25–30 и 50–55, т.е. граничными областями $V_{мл} \cap V_{ср} \neq \emptyset$ и $V_{ср} \cap V_{ст} \neq \emptyset$.

Нечеткое множество A состоит из неопределенного числа элементов $x \in X$, где X – универсальное множество. Для каждого элемента множества A задается функция принадлежности $\mu_A(x)$, $0 < \mu_A(x) < 1$, выражающая степень принадлежности его множеству A . Подмножество $S \subset X$, $S = \{x \in X | \mu_A(x) > 0\}$, содержащее элементы с $\mu_A(x) > 0$, называется носителем нечеткого множества.

Пример. Нечеткое понятие «возраст юноши» характеризуется носителем $13 \leq S \leq 30$, на котором определено нечеткое множество A с функцией принадлежности $\mu_A(x)$.



Здесь понятию «возраст юноши» соответствует интервал от 13 до 30 лет, причем полное соответствие наблюдается для возраста от 15 до 25 лет.

Нечеткое отношение R представляет собой подмножество декартова произведения $X \times X$, в котором каждая пара $x_i R x_j$, $x_i, x_j \in X$ охарактеризована численно функцией принадлежности $\mu_R(x_i, x_j)$. Примерами нечетких отношений являются «примерно равен», «значительно больше», «предпочтительнее» и т.д.

У объектов гуманитарных исследований, как правило, связь между качественными и количественными характеристиками выражена неявно, и в этом сложность и специфика количественных исследований в гуманитарных науках. Вместе с тем нечеткие отношения позволяют моделировать такую связь вне зависимости от ее вида: линейную, нелинейную и т.д. Так, с помощью нечетких отношений могут быть решены задачи моделирования причинно-следственных связей при изучении различных процессов в гуманитарных исследованиях. Причина и следствие (вход и выход системы, процесса) могут быть описаны нечеткими множествами A и B соответственно.

Формализовать имеющуюся причинно-следственную зависимость ($A_1 \rightarrow B_1$, $A_2 \rightarrow B_2$, ..., $A_n \rightarrow B_n$) можно с использованием нечеткого отно-

шения R на основе правила композиции нечетких выводов

$$B = A \circ R \quad (1)$$

Уравнения нечетких отношений (1) применяются как для получения восходящих выводов $A_{n+1} \rightarrow B_{n+1}$ (от предпосылки к заключению – задачи прогнозирования), так и для получения нисходящих выводов $B_{n+1} \rightarrow A_{n+1}$ (от заключения к предпосылке – например, при оценке (уточнении) причин тех или иных социальных явлений по нечеткой или неполной информации).

Теория нечетких отношений находит также приложения в задачах, в которых традиционно применяется теория обычных (неразмытых, четких) отношений. Как правило, аппарат теории четких отношений используется при качественном анализе взаимосвязей между объектами исследуемой системы, когда взаимосвязи носят дихотомический характер и могут быть проинтерпретированы в терминах «связь присутствует», «связь отсутствует» либо когда методы количественного анализа взаимосвязей по каким-либо причинам неприменимы и взаимосвязи искусственно приводятся к дихотомическому виду. Например, когда величина связи между объектами принимает значения из ранговой шкалы, выбор порога силы связи позволяет преобразовать связь к требуемому виду. Однако подобный подход, позволяя проводить качественный анализ систем, либо приводит к потере информации о силе связей между объектами, либо требует проведения вычислений при разных порогах. Этого недостатка лишены методы анализа данных, основанные на теории нечетких отношений, которые позволяют проводить качественный анализ систем с учетом различия в силе связей между объектами системы.

Для гуманитарных исследований обычно характерно наличие большого числа исследуемых факторов. Даже когда классические математические методы теоретически могут быть применены для решения некоторой проблемы, часто их практическое использование оказывается затруднительным именно из-за большой размерности задачи.

К таким задачам, прежде всего, относятся задачи классификации в гуманитарных исследованиях. Классификация осуществляется, как правило, по многим параметрам. Это задачи типологизации в социологии, обработки анкет и т.п.

При этом названия групп классификации, например при классификации библиотек – «сверхгиганты», «гиганты», «сверхкрупные библиотеки», – достаточно условны. Эти понятия не имеют четких границ и ассоциируются одновременно с объемными фондами, большим читательским контингентом, многочисленным и квалифицированным обслуживающим персоналом, разнообразными

формами библиотечного обслуживания и многим другим.

Наиболее существенный недостаток традиционных моделей классификации связан с лежащими в их основе бинарной логикой принадлежности и принципом однозначного (четкого) разделения объектов на классы.

Любая классификация представляет собой структуру, порожденную отношением подобия, которое в общем случае представляет собой не-транзитивное бинарное отношение. Другими словами, если A подобно B и B подобно C , далеко не всегда следует, что A подобно C . Нечеткая логика позволяет учитывать это обстоятельство, тогда как структура, обусловленная бинарной логикой принадлежности обычных (не нечетких) множеств, определяет транзитивное отношение между парами представителей классов следующим образом: из « x принадлежит тому же классу, что и y » и « y принадлежит тому же классу, что и z » всегда следует, что « x принадлежит тому же классу, что и z », $\forall x, y, z \in X$.

Существенное преимущество нечеткого подхода состоит также в том, что в рамках многозначной логики осмысленные решения находятся для более широкого класса проблем, нежели при четкой постановке [8]. Действительно, если в обычных классификационных моделях объект $x_i \in X$ может принадлежать либо не принадлежать классу F_j , т.е. соответствующая функция принадлежности μ_{ij} может принимать только два значения – 0 или 1, то в моделях нечеткой классификации значения функции μ_{ij} непрерывно покрывают весь единичный интервал, т.е. $\mu_{ij} \in [0, 1]$.

Типичный пример задачи нечеткой классификации – автоматическая классификация библиографических записей при создании электронного каталога [9]. Каждая библиографическая запись должна быть отнесена к определенному классу (предметной рубрике). Классы не имеют четко выраженных границ, в связи с чем принадлежность записи тому или иному классу часто не может быть определена однозначно. Кроме того, принадлежность записи тому или иному классу определяется в ходе автоматической классификации на основе сравнения с имеющимися представителями кластеров, при этом отношение сходства является нечетким. Таким образом, появляется возможность организации поиска в массивах информации по нечетким признакам.

Для решения данной задачи могут быть использованы следующие подходы.

Первый подход основан на определении принадлежности записи определенному классу по максимальному значению степени включения [10].

Пусть $a = \{a_1, a_2, \dots, a_n\}$ – библиографическая запись, где a_i – термы, образующие запись ($i = 1,$

$\dots, n$), $\{L_1, L_2, \dots, L_m\}$ – классы (предметные рубрики), к одному или нескольким из которых следует отнести библиографическую запись. Пусть $L_j \neq \emptyset$ ($j = 1, \dots, m$). На основе имеющихся статистических данных по записям, содержащимся в классах L_j строим матрицу $M = \{\mu_{ij}\}$, элементами которой являются μ_{ij} – степени принадлежности элемента a_i (или совокупности элементов) классу L_j ($i = 1, \dots, n$). Так как информативность термов (или их совокупности) различна, μ_{ij} домножаем на коэффициент значимости k_i ($0 < k_i < 1$). Коэффициенты значимости определяем экспертным путем.

Таким образом, имеем совокупность классов $\{L_j\}$, каждый из которых можно представить множеством $\tilde{L}_j$, и запись a , заданную как нечеткое множество $\tilde{a}$ векторами $\mu_j = \{k_1\mu_{1j}, k_2\mu_{2j}, \dots, k_n\mu_{nj}\}$. На основе этой информации определяем степень включения $\tilde{a}$ в $\tilde{L}_j$:

$$v_j = v(\tilde{a}, \tilde{L}_j) = \bigwedge_{a_i \in a} (1 - \mu_a(a_i)) \vee \mu_{L_j}(a_i) \quad (2)$$

По максимальному значению v_j ($j = 1, \dots, m$) определим класс, которому принадлежит запись a .

Для определения принадлежности записи одному из заданных классов могут быть использованы и другие подходы:

- классификация записей на основе построения нечеткого отношения сходств данной и эталонной записи;
- классификация записей с использованием ассоциативных сетей.

Для построения эталонной записи строится дерево классов. Для каждого класса определяется эталонная запись, характеризующая класс. Естественно, эта запись должна быть нечеткой. Затем строится нечеткое отношение сходства эталонной и заданной записей. По степени сходства с эталонной записью определяется принадлежность записи определенному классу.

При построении ассоциативной сети сначала идентифицируются классы, характеризующиеся некоторыми или всеми признаками, имеющимися у записи. После этого в полученной совокупности классов идентифицируется наиболее значимый класс.

Одно из замечательных свойств ассоциативной сети заключается в том, что при классификации записей она явно приближается к модели человеческого мыслительного процесса.

Перечисленный круг задач и примеры их решения с использованием теории нечетких множеств позволяют сделать вывод о перспективности использования нечеткой логики в гуманитарных исследованиях.

В философском плане теория нечетких множеств примечательна тем, что открывает новый

подход к решению проблемы абстракции и образования понятий, обладающих богатством всевозможных оттенков. В области психологии – это моделирование свойств целостности, диффузности психических образов и представлений, гибкости мышления, многозначности элементов языка на всех уровнях коммуникации. В области лингвистики – это моделирование смысла предложений и текстов с помощью распределения возможностей, описываемых функциями принадлежности.

В заключение отметим, что строго формализованный количественный анализ имеет свои пределы [11], за которыми могут быть утрачены качество, глубина и полнота осмысления действительности. Поэтому в гуманитарных исследованиях необходимо использовать «гибкие» методы, одним из которых и является нечеткая логика.

Список литературы

1. Zadeh, L. A. Fuzzy sets // Information and Control. – 1965. – Vol. 8. – P. 338–345.
2. Ядов, В. А. Стратегия социологического исследования / В. А. Ядов. – М. : Добросвет, 1998. – 596 с.
3. Толстова, Ю. Н. Измерение в социологии / Ю. Н. Толстова. – М. : Добросвет, 1997. – 248 с.
4. Saaty, T. L. Exploring the interface between hierarchies, multiple objectives and fuzzy sets // Fuzzy Sets and Systems. – 1978. – Vol. 1. – P. 57–69.
5. Беллман, Р. Принятие решений в расплывчатых условиях / Р. Беллман, Л. Заде // Вопросы анализа и процедуры принятия решений : сб. переводов. – М. : Мир, 1976. – С. 172–215.
6. Крумберг, О. А. Теория психологической возможности для моделирования выбора в условиях неопределенности // Методы принятия решений в условиях неопределенности. – Рига, 1980. – С. 47–52.
7. Нечеткие множества в моделях управления и искусственного интеллекта / под ред. Д. А. Поспелова. – М. : Наука, 1986. – 312 с.
8. Вершинина, Л. П. Нечеткий классификационный алгоритм распознавания образов // Медико-экологические информационные технологии-98 : материалы междунар. техн. конф. – Курск, 1998. – С. 209–211.
9. Вершинин, М. И. К вопросу о ретроспективной конверсии карточных каталогов // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества : материалы конф. – Судак, 1999. – С. 588–589.
10. Вершинин, М. И. Программа VerIsis / М. И. Вершинин, Л. П. Вершинина. – № Гос. ФАП 50980000058. – 1998.
11. Шубкин, В. Н. Пределы // Новый мир. – 1978. – № 2. – С. 127–135.

Материал поступил в редакцию 08.10.2007 г.

Сведения об авторах: Вершинин Михаил Иосифович – кандидат педагогических наук,
помощник ректора по информатизации,
Вершинина Лилия Павловна – доктор технических наук,
профессор, заведующая кафедрой